

Title	睡眠するロボットの開発
Author(s)	浅香, 智輝
Citation	平成30年度学部学生による自主研究奨励事業研究成果報告書
Issue Date	2019-04
oaire:version	VoR
URL	https://hdl.handle.net/11094/71959
rights	
Note	

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

平成30年度学部学生による自主研究奨励事業研究成果報告書

ふりがな 氏 名	あさか ともき 浅香 智輝	学部 学科	基礎工学部シス テム科学科	学年	2 年
ふりがな 共 同 研究者氏名		学部 学科		学年	年
					年
					年
アドバイザー教員 氏名	まえ やすし 前 泰志	所属	基礎工学研究科 システム創成専攻		
研究課題名	睡眠するロボットの開発				
研究成果の概要	研究目的、研究計画、研究方法、研究経過、研究成果等について記述すること。必要に応じて用紙を追加してもよい。(先行する研究を引用する場合は、「阪大生のためのアカデミックライティング入門」に従い、盗作剽窃にならないように引用部分を明示し文末に参考文献リストをつけること。)				

1. はじめに

睡眠は、疲労回復等の恒常性維持の他に、記憶や学習に対して大きな役割を果たしており、人間は、覚醒と睡眠を繰り返すことで、高度な学習を行っている。ロボットの機械学習において、実機での学習を覚醒、シミュレーション上での学習を睡眠とみなすことで、このサイクルを実用できないかと考えた。

本研究の意義は、アクチュエータへの負荷を軽減しつつ、学習の効率化を図ることである。機械学習によって、ロボットに歩行等の運動を学習させる際、カメラの FPS やロボット転倒時の復帰作業等がボトルネックとなり、長い学習時間が必要になる。その間にロボット本体が破損する恐れがあるため、シミュレーション上で学習を行い、最終的な学習データを実機で利用することがある。しかし、ロボットの自由度が高い場合や、環境やタスクが複雑な場合、シミュレーションと実機環境の差が、学習結果に一定の影響を及ぼすといえる。そのため、正確なシミュレーション環境が必要になり、機械学習のハイパーパラメータに加えて、調整すべきパラメータが大量になる。

上記の問題に対し、シミュレーションで転倒しない運動を学習し、実機でシミュレーション環境との差を補正する学習を行うことで、ある程度効率化が図れるのではないかと考えた。

2. 研究目的

本研究の目的は、実機とシミュレーションを併用し、多脚ロボットにリーチング学習を行わせる事である。具体的な目標は、シミュレーションのみで学習を行った場合と、実機とシミュレーションを併用した場合を設定し、最終的な実機環境でのテストで、後者でより良い結果を得ることである。

3. 装置及び環境

本研究では、多脚ロボットに、18 自由度を持つ近藤科学株式会社の KXR-L6 を使用した (図1)。同社の RCB-4mini コントロールボードに、同社の Bluetooth モジュールの KBT-1 を搭載し、各サーボを任意に遠隔操作できるようにした。

実機の位置は、ロボットにマーカーを貼りつけ、これを画像認識によってトラッキングした(図2)。これは、OpenCV を用いて、Web カメラから取得した RGB 色空間画像を、HSV 色空間画像に変換後、指定した HSV 範囲について2値化し、プロブ解析をすることで行った。

図2において、2つの十字は、それぞれマーカーの重心座標、ターゲットの座標を示している。

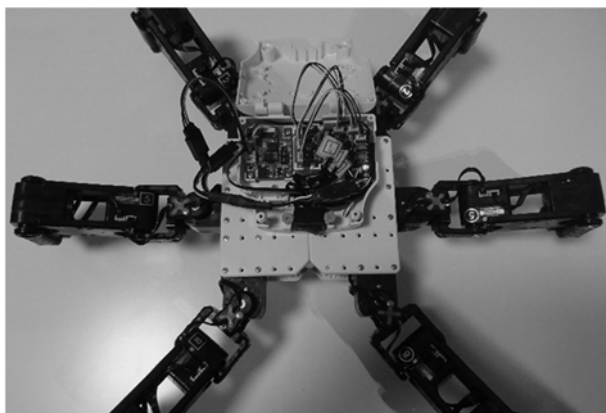


図1 多脚ロボット(KXR-L6)

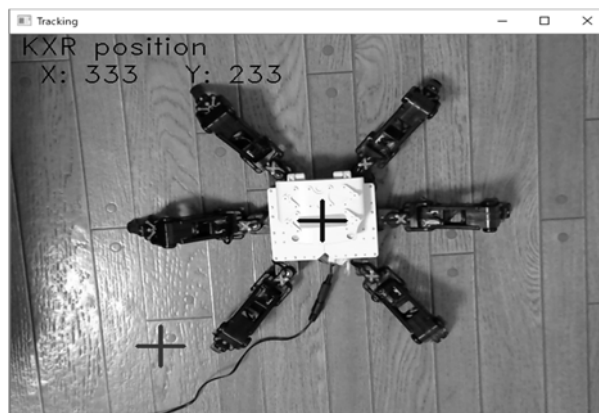


図2 トラッキング

シミュレーション環境には、Unity の ML-Agents 環境[1]を使用し、Python によって、TensorFlow のマルチプロセス機能を利用し、学習アルゴリズムは OpenAI の PPO を用いて実験を行った。Unity のモデルは、ロボットの主要パーツのみを再現し、足を2つのグループに分け、トライポッド歩容するように制限を付けた。また、サーボ角についても、互いに干渉しないように制限を付けた。マルチプロセスで並列処理するエージェントは9体とした(図3)。

タスクとしてリーチング学習を設定し、図3における各ステージ上の立方体をターゲットとした。エージェントがターゲットに到達すると、報酬が与えられ、ターゲットがステージ上にランダムに再配置される。エージェントの初期位置は、各ステージの中央である。図3において、中央のエージェントが、実機と対応しており、実機での学習時は、サーボ角やターゲットの位置情報が共有される。

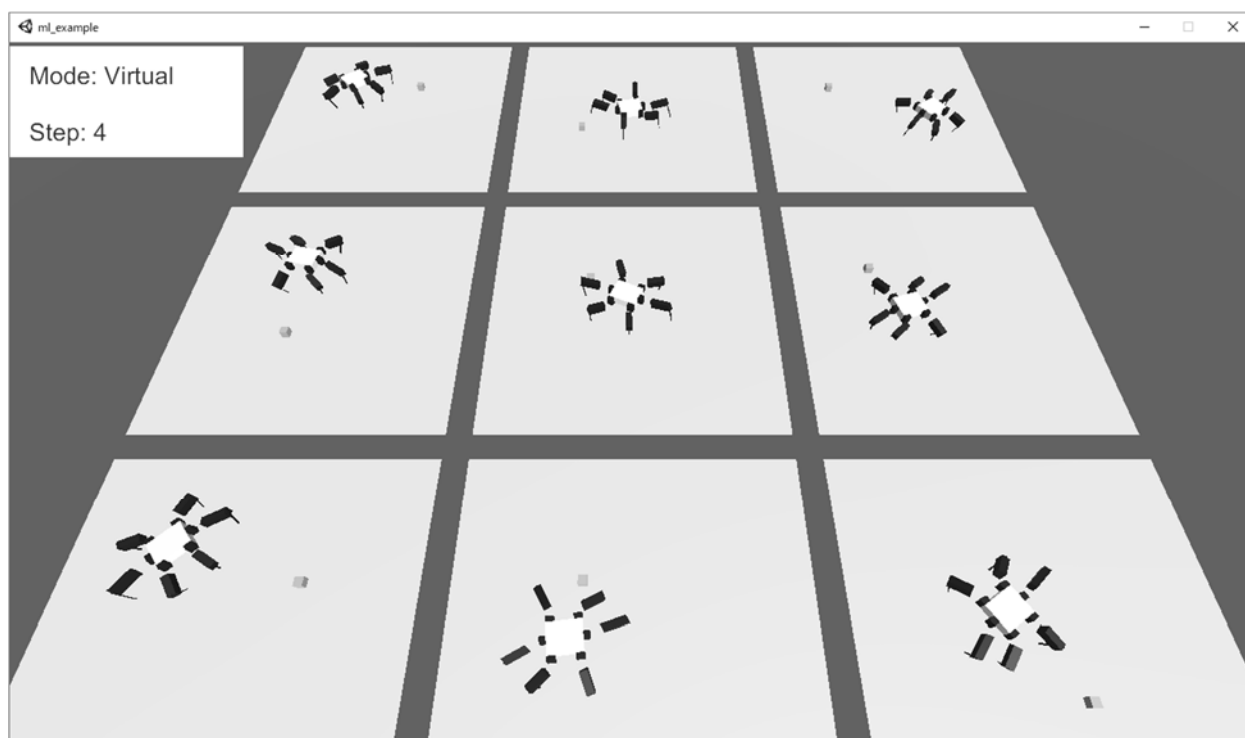


図3 Unity のシミュレーション環境

エージェントの報酬は、ターゲットとの距離の逆数を評価値としており、他にも、胴体部の傾きが一定角度内に収まっている時に正の報酬を与え、外れた場合は負の報酬を与えた。また、胴体が接地しそうな場合や、ステージ外に出そうになった場合にも、負の報酬を与えるようにした。ただし、前者の場合は、その場でサーボ角をホームポジションに強制的に直し、後者の場合は、ステージ中央から再スタートさせた。サーボのホームポジションは、KXR-L6 のサンプルモーションと同じ角度とした。実機の場合も、同様の報酬を与え、ステージ外に出た時のみ復帰作業を行った。各ステップにおいて、常に負の報酬を与えることで、より早くタスクを完了するようにした。

4. 実験

実機とシミュレーションの学習形態及び、期待する結果を、表 1 に示す。実機のみでの学習は、上記のように、長い学習時間がかかる事に加えて、その間に樹脂サーボギヤ等のパーツが破損する可能性があったため、本研究では実施しなかった。

リーチング完了時に、ステップを 1 増加させ、9 体のエージェントのステップ総数を用いて、実機又はシミュレーションで学習させるか選択するようにした。ステップ総数が 100 回以上になった場合に、実機テストを行った。実機での学習割合は、ステップ総数の 5%とした。

表 1 学習形態

学習形態	実機テストで期待する結果
シミュレーションのみ	リーチング不可または、リーチングに長時間要する
実機及びシミュレーション	シミュレーションのみの場合よりも、効率的にリーチングを行う

5. 結果・結論

シミュレーションにおける、ステップ数あたりの、各エージェントのタスク継続時間の平均値の推移を、図 4 に示す。各エージェントのタスク継続時間は、タスク完了時にそれぞれリセットされ、Unity の経過フレーム数をカウントする。図 4 において、学習開始時は、フレーム数が蓄積されていないため、偶然タスク完了した際に、小さい値が現れている。学習が進むにつれて、各エージェントのリーチング継続時間が蓄積されていき、ステップ数が 50 回となった周辺で最大となっている。その後、ステップ毎にリーチング継続時間が減衰しているため、Unity 上で学習が進んだことが分かる。

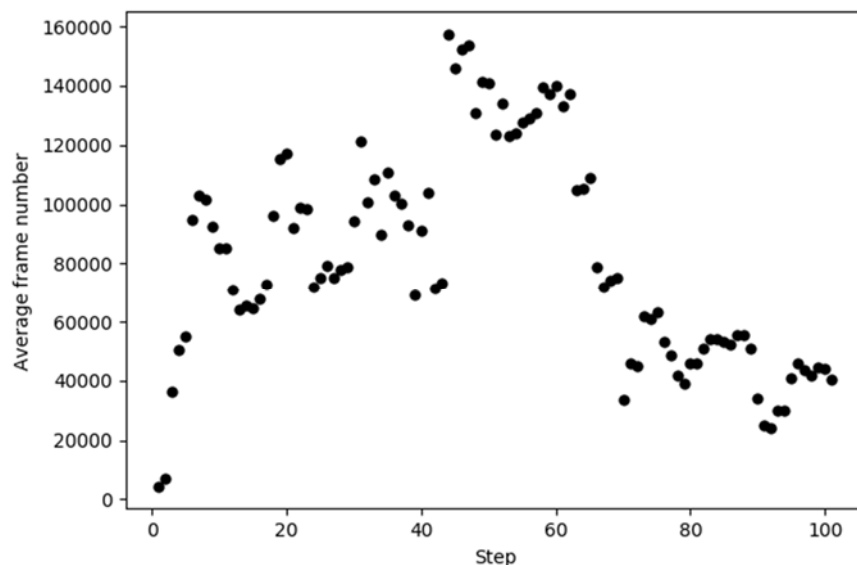


図 4 リーチング継続時間（シミュレーション内）

表 1 の学習形態で機械学習を行い、実機テストで得られた結果を、表 2 に示す。シミュレーションのみの場合は、実機テストにおいて、実時間 15 分 38 秒で画面外にロストし、リーチングを完了することができなかった。

実機及びシミュレーションの場合は、実機テストにおいて、実時間 1 分 5 秒でリーチングを完了した。実機での学習中は、1 回目と 4 回目でロボットが画面外にロストしてしまい、他の場合ではリーチングを完了した。

実機とシミュレーション環境の差について、実機ではフローリング上で実験を行ったため、摩擦力による影響が大きく目立った。実機テスト時に、Unity 上で対応しているエージェントとの位置や向きに差が生じ、画面外にロストしてしまったのだと考えられる。歩行する度に、地面や床などの摩擦係数を計測することは現実的でないため、このような要因に対しては、本研究の手法が有効であると考えられる。

以上の結果から、実機とシミュレーションを併用した場合の方が、実機テストにおいて、良い成績を出した。そのため、本研究の目的について、一定の成果を得ることができた。しかし、データが学習過渡期のものであり、収束した後のものではないため、やや信頼性に欠ける結果だと言える。

表 2 実機テストでの結果

学習形態	実機テストで得られた結果
シミュレーションのみ	実時間 15 分 38 秒で画面外にロストし、リーチング不可
実機及びシミュレーション	実時間 1 分 5 秒でリーチング完了

6. 今後の課題

本研究では、機械学習時に、ハイパーパラメータの調整に十分な時間を割くことができず、他にも、Unity で使用したモデルの運動制御や Bluetooth の安定動作、サーボに加わる電圧や電流の管理などの点も妥協してしまった。実機での学習時、Bluetooth 通信と OpenCV での画像処理をするため、Unity のフレームレートが極端に遅くなってしまった。そのため、ロボットにオンボード処理をさせるべきであったが、既存のロボットはホビー向けが多いため、サーボトルクやスペースの問題がある。Unity では CAD のデータを取り込んで使用できるため、ロボットの設計から行い、精度の高いシミュレーション環境を実装し直して、より信頼できる結果を示したい。

7. 謝辞

本研究を行うにあたり、基礎工学研究科のシステム創成専攻の前泰志准教授から、様々なご指導を頂きました。お忙しい中、お時間を割いて頂き、多くの事を学ばせて頂きました。また、このような機会を提供して下さい、自主研究奨励事業を支えて下さる皆様に、この場をお借りして感謝申し上げます。

参考文献

[1] xiaomaogy 「Unity Machine Learning Agents Toolkit」

<https://github.com/Unity-Technologies/ml-agents> (2018/12/13 最終閲覧)